



Memory Solutions for Diverging Applications

A New Perspective on Performance

Terry Lee, Micron Technology, Inc.

Executive Director

Advanced Technology and Strategic Marketing

tleee@micron.com



Agenda

- **Changing perspective of memory performance and development**
- **Performance trends and solutions in:**
 - Handheld/wireless
 - Networking
 - Graphics/entertainment
 - Servers
 - Laptops
 - Desktop PCs
- **Summary**

Memory Performance – A Historical Perspective

- ▣ **One memory technology served nearly all applications**
 - ▣ Customers optimized their system designs to use the highest-volume memory device
 - ▣ Performance was improved over time by increasing clock rates
- ▣ **Slow data rates allowed system implementation using simple digital design techniques**
- ▣ **Customer engineering departments provided the majority of expertise needed for system implementation**
- ▣ **Supplier technical support focused on expertise in device features and protocol**

Memory Performance – Today's Perspective

- ▣ **Increasing emphasis on latency, bandwidth, and power dissipation – mutually exclusive priorities**
 - ▣ Segmentation of the memory market
- ▣ **Increased priority on innovation**
 - ▣ Decreased priority to use standard memory device in some applications
- ▣ **More complex memory requirements are demanding more engineering resources from suppliers and customers**
- ▣ **Ability to innovate and drive new technology is limited to the strongest players**

Priorities for Memory Applications



Desktop
Bandwidth
Volume/cost

Consumer/Entertainment

Very high bandwidth
Thermal/cost



Notebook

Small form factor
Low power
Thermal
Volume/cost



Server

Density
Low power
Thermal
Desktop leverage



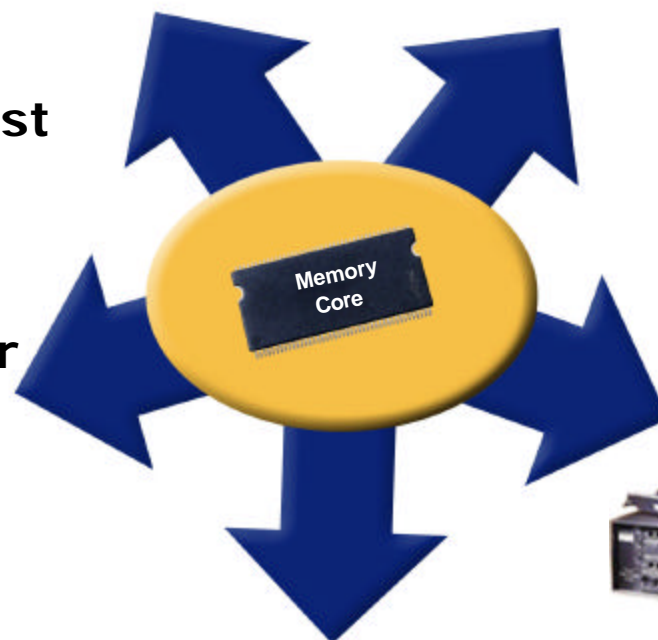
Network

Low latency
High bandwidth
Low power



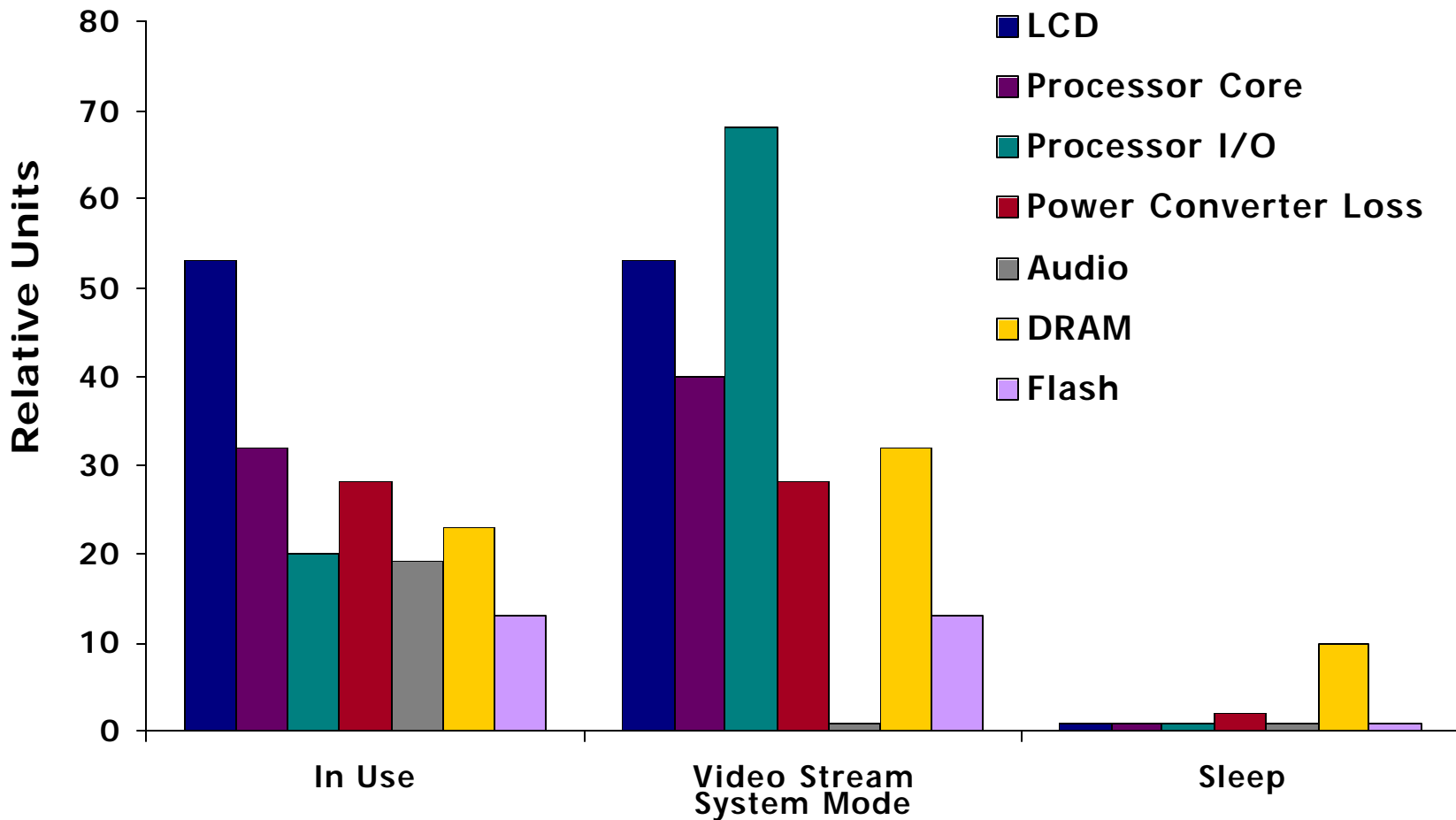
Handheld/Wireless

Small packaging
Very low power
Moderate cost



From Vision to Reality

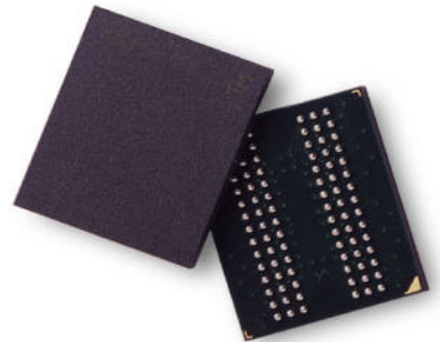
Handheld Power Usage



From Vision to Reality

Handheld Solution – Mobile SDRAM

- **Lower operating voltage to reduce power dissipation in all operating modes**
- **Reduced standby (sleep) power by adding new operating modes**
 - Partial array self refresh
 - Temperature compensated refresh
- **Overall power reduction**
 - Operating >40%
 - Standby/self refresh >80%, but even greater when partial array mode is used!
- **Small form factor packaging**
 - FBGA and bare die

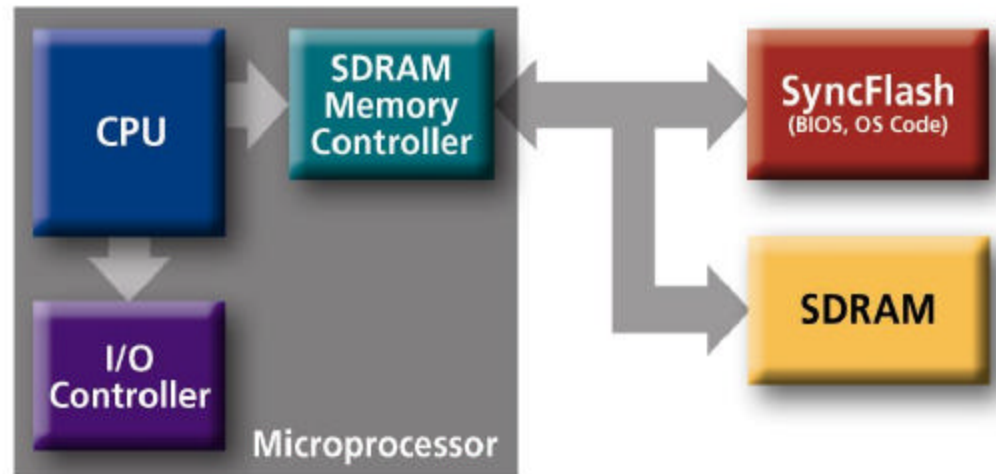


Instant-On for Handhelds

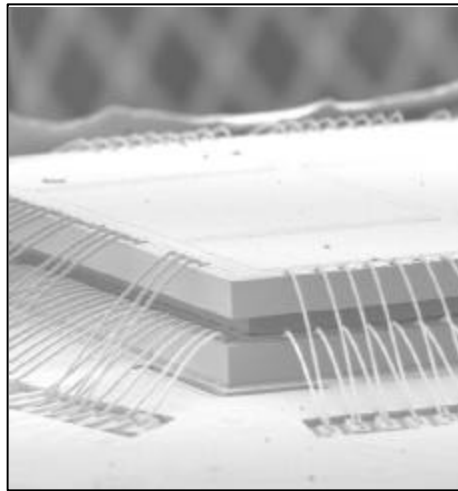
□ **Solution: SyncFlash®**

- NOR Flash with SDRAM-compatible interface
- Resides on SDRAM bus
- Direct boot from SyncFlash eliminates boot ROM
- Execute-in-place at 133 MHz
- System ready in 1-2 seconds!

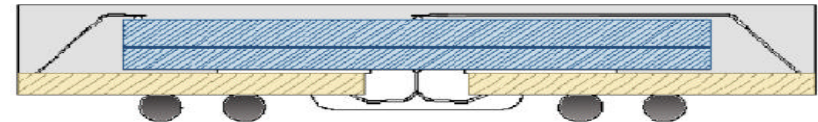
Native SyncFlash Controller



Handheld Solution - Packaging Options



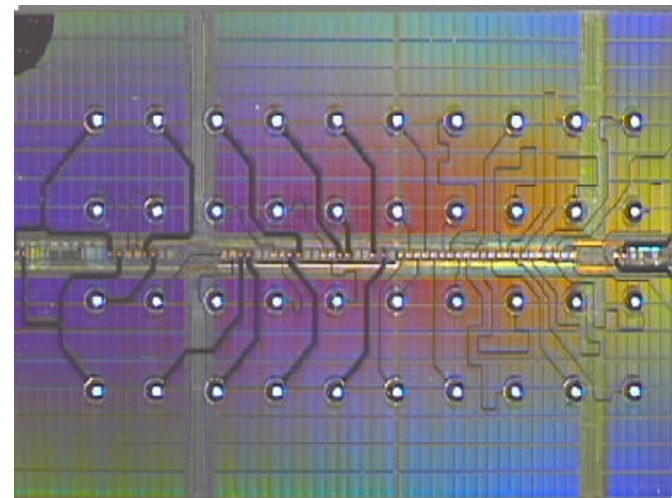
Same-Size Stacked Die



2BOC (FBGA)

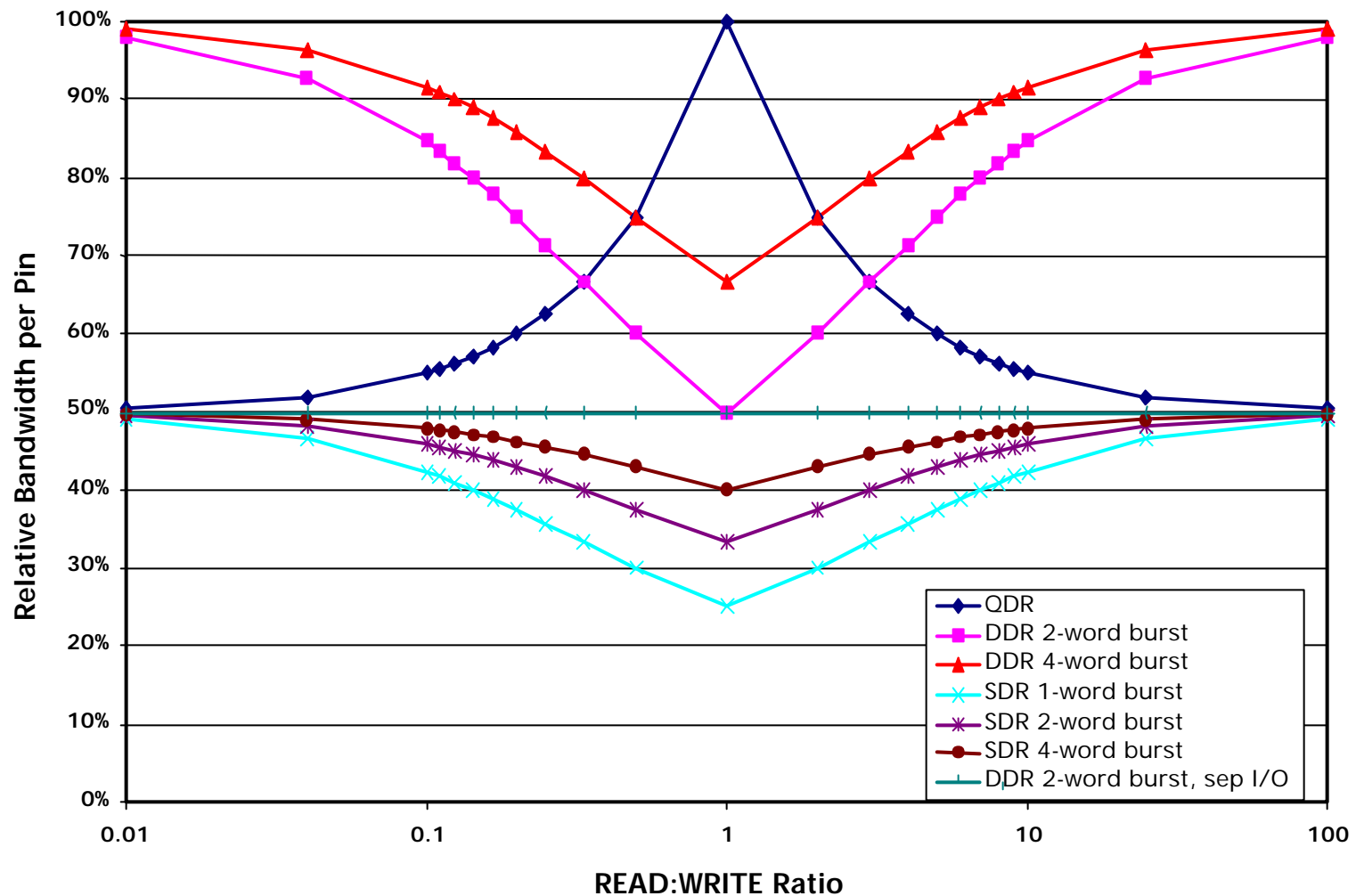


2SOP



Wafer Level Re-distribution Layer

Network Performance



Network Solution – New DRAM Architectures

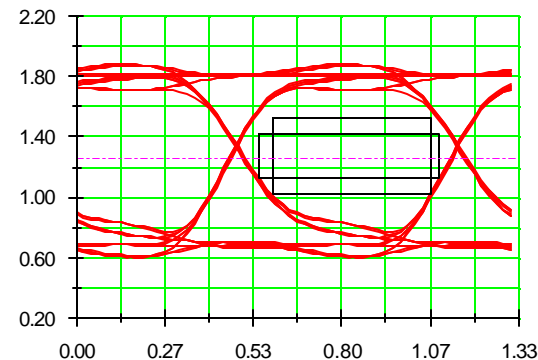
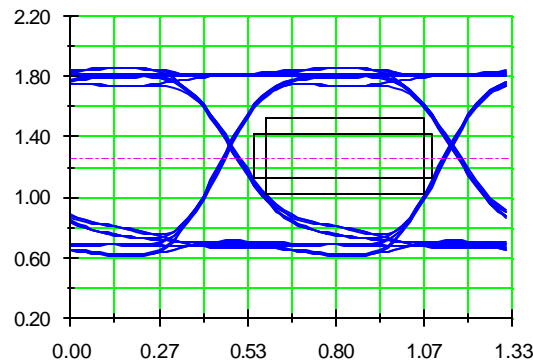
- ❑ **SRAM devices used in networking designs are high cost per megabit**
- ❑ **New DRAM architectures: RLDRAM™ and FCRAM**
- ❑ **RLDRAM (Reduced Latency DRAM)**
 - ❑ Low random bank cycle or t_{RC} for data bus efficiency
 - ❑ High bandwidth (800 Mb/s)
 - ❑ Low core and I/O voltages for power concerns
 - ❑ Minimized package size for space considerations
- ❑ **Future trends**
 - ❑ Higher bandwidth, lower latency, bare die, and stacked die packaging

Network Solution - Ternary CAMs

- **What do TCAMs do ?**
 - Search all memory locations at exactly the same time
 - Provide network line cards with lightning-quick searches in hardware vs. slower software searches of conventional memory
- **Binary CAMs store only 0s and 1s – “Exact Match” searches**
- **TCAMs store 0s, 1s, and Xs (don’t care)**
 - Group addresses into a single table entry
 - “Longest Prefix Match” searches

Graphics/Entertainment

- **Challenge: very high bandwidth, relatively low cost**
- **Solution: GDDR3**
 - x32 organization
 - New signaling interface
 - On-die termination
 - Flip-chip packaging
 - Ultra-high bandwidth – 1.0 Gb/s to 1.5 Gb/s



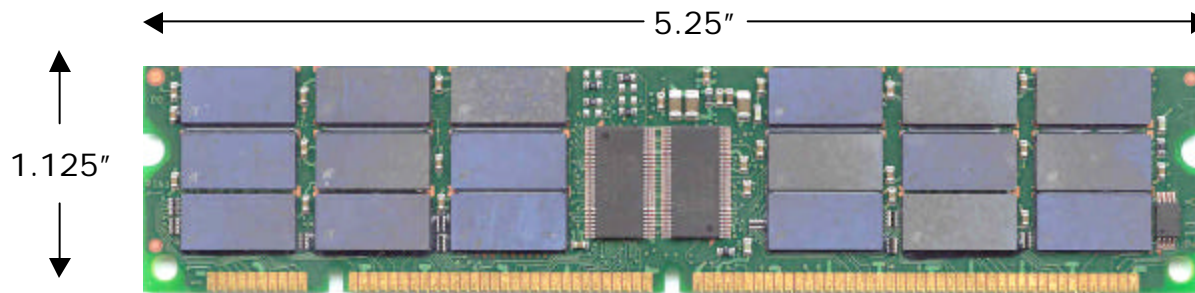
GDDR3 at 1.5 Gb/s

Servers

□ Challenge:

- Lower standby and operating power due to number of devices in system
 - Memory design challenge
- High data rates require lower inductance packaging; low-profile DIMMs required for rack servers
- How to build stacked modules?

□ Solution: **2BOC and flip-chip**



1GB Flip Chip Stacked Module

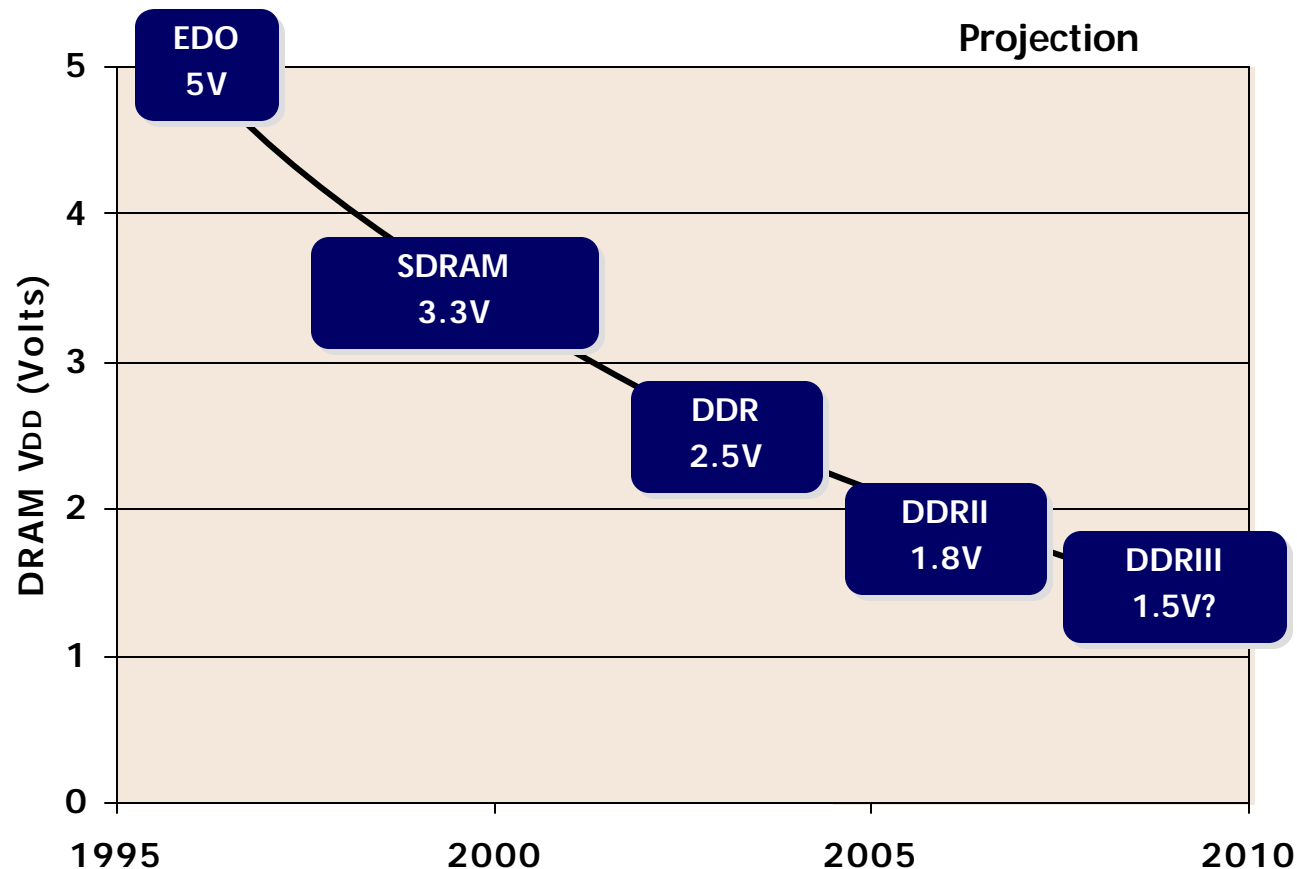
Laptops

- **Challenge:** Low power, small form factor, low price
- **Solution:**
 - DDRII
 - SODIMMs populated with FBGA components
 - 1.8V operating and I/O voltages
 - Power dissipation at 400 Mb/s is reduced by 20% over DDR266
 - On-die termination reduces system components and board real estate
 - Data rates increase to 400 Mb/s and 533 Mb/s



FBGA-Based, Small-Outline DIMM

Main Memory Voltages



- VDD is being reduced with each new technology
- I/O voltages are also being reduced
- Each decrease of VDD allows significant power reductions

Desktop Performance

□ **Challenge:**

- Always increase the bandwidth
- Never increase the cost
- Power dissipation is becoming a challenge

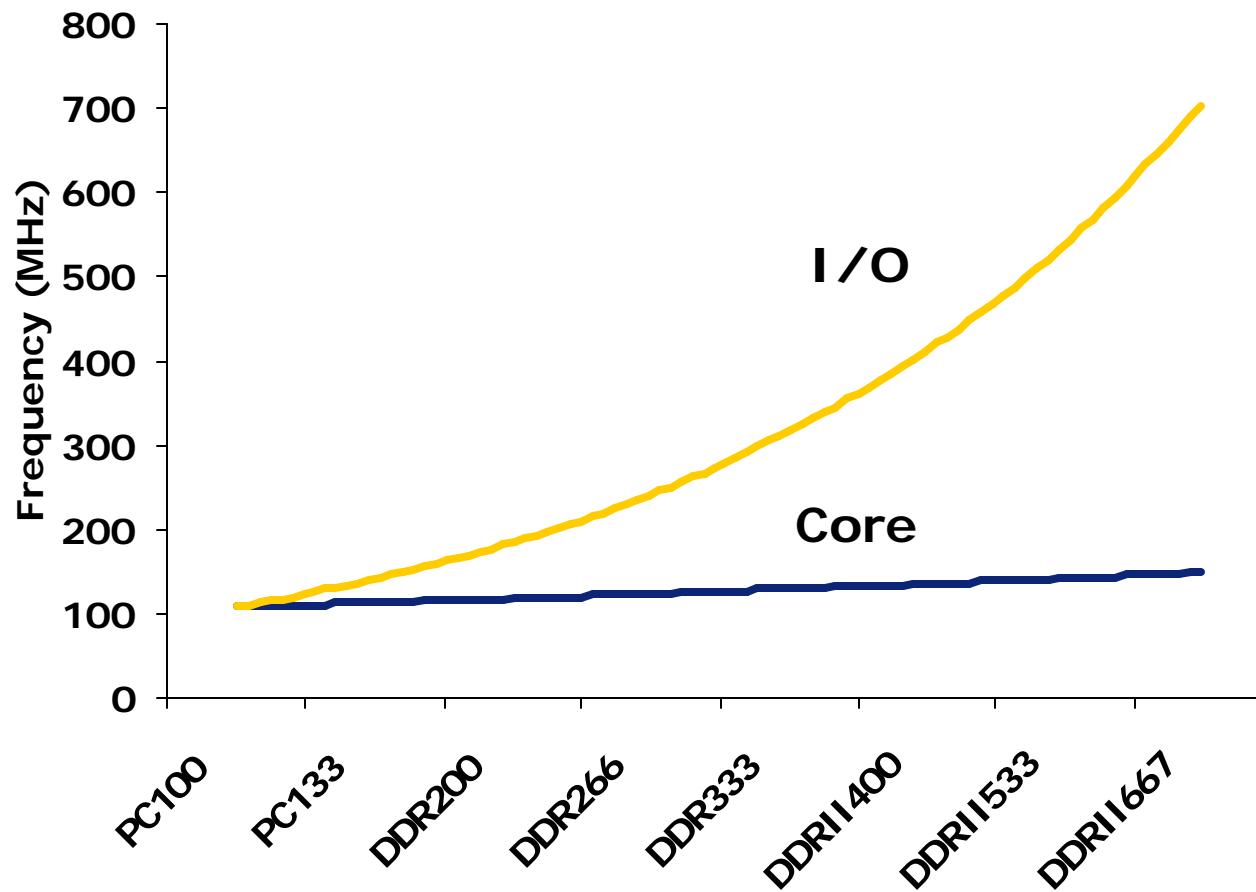
□ **Solution:**

- Architecture must migrate at each technology transition
- Lower operating voltage
- Improve signaling
- Design of memory channel requires intensive engineering
 - Collaboration of system designer and memory supplier is important

Techniques For Increasing Bandwidth

- **Burst modes (SDR, DDRI, DDRII, etc.)**
 - Improves command saturation; tailored to cache lines
 - Inexpensive to implement
- **Multiple strobe edges**
 - Basis of double data rate
- **Prefetching (2n, 4n, etc.)**
 - Pipelining data; minimum WRITE burst
 - Fairly inexpensive die cost; saves power in core
- **Increased core speed**
 - Column-to-column cycle time
 - Gradual improvements; total power reduction higher priority
- **Improved termination techniques**
 - On-die termination for DDRII

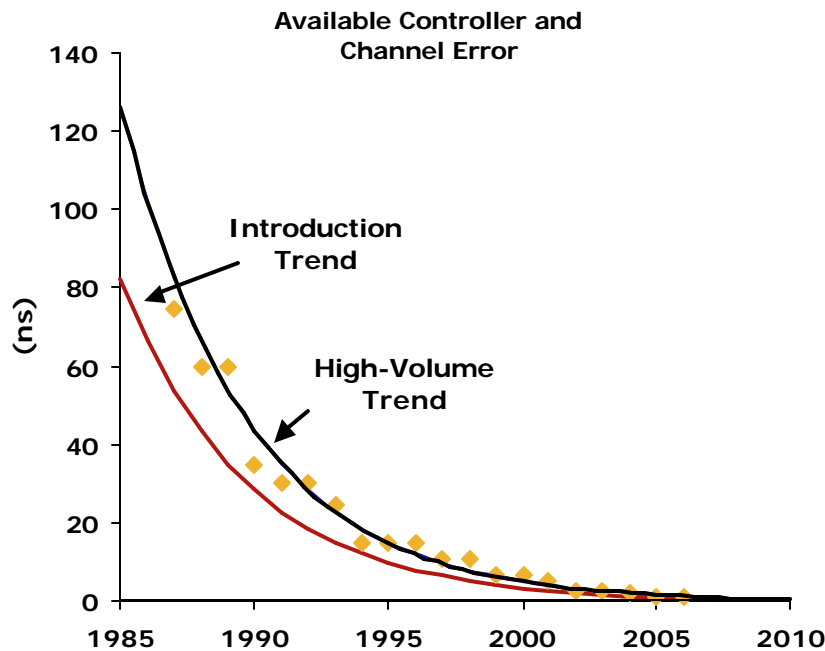
Bandwidth vs. Core Speed Trend



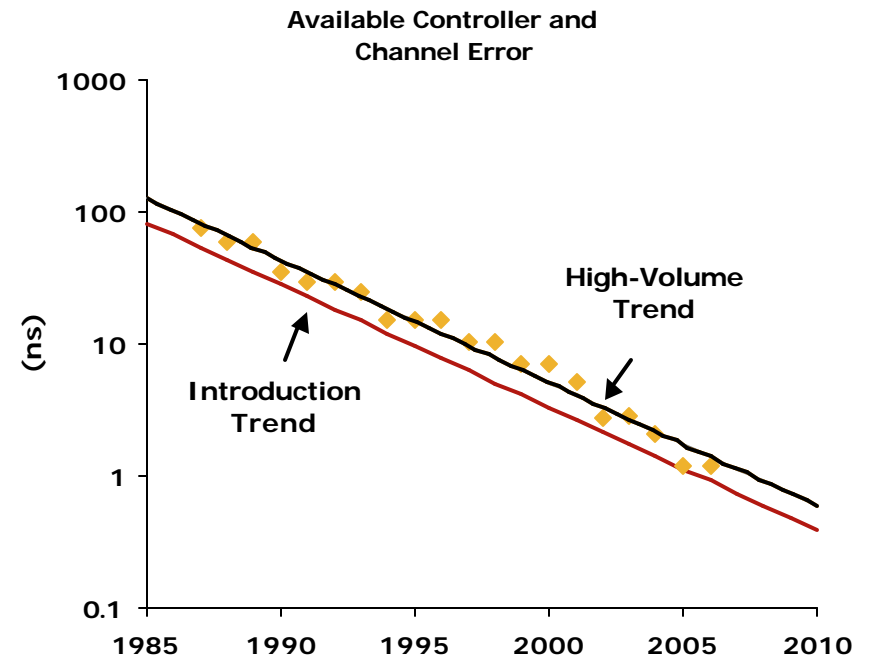
From Vision to Reality

Faster Data Rates Require More Engineering of Signaling and Physical Design

From Vision to Reality



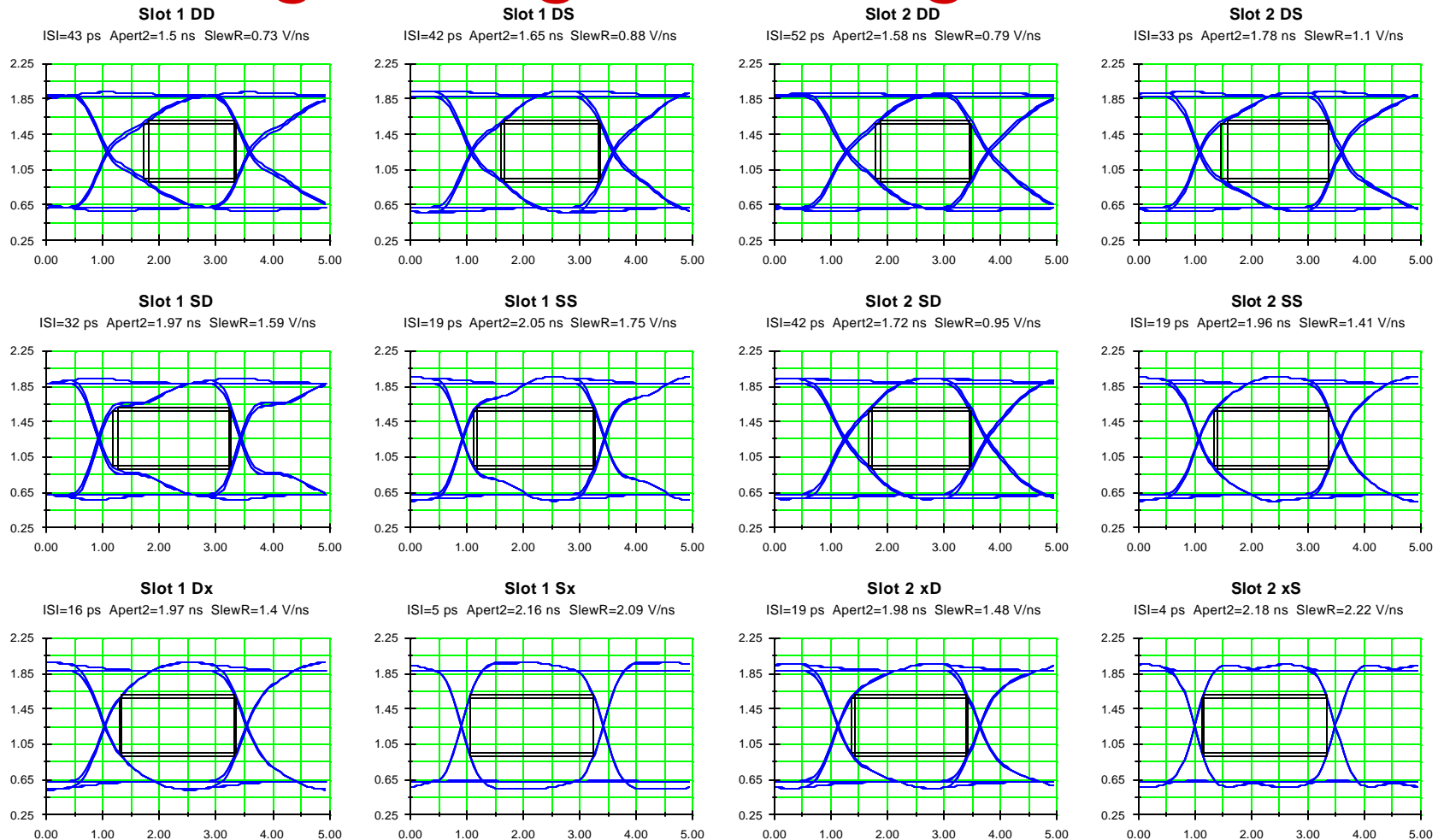
Linear Scale



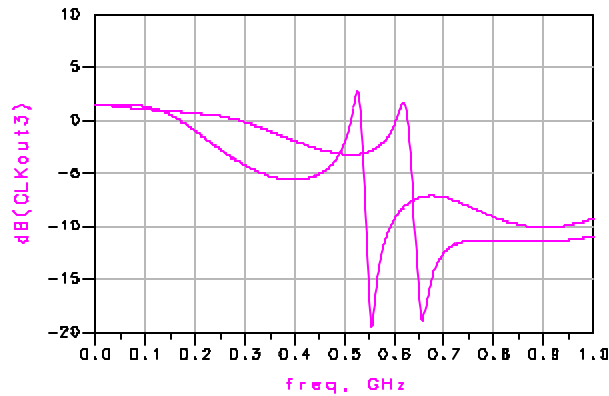
Logarithmic Scale

Multidrop Desktop Environment is a Signaling Challenge

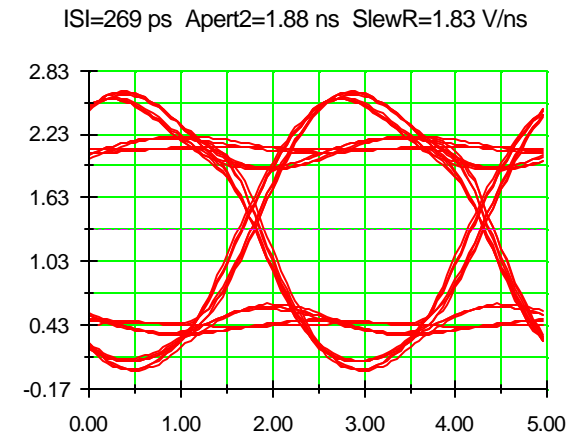
From Vision to Reality



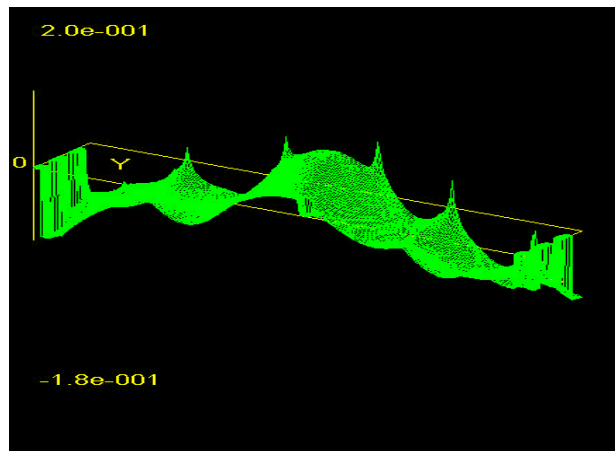
Platform Analysis



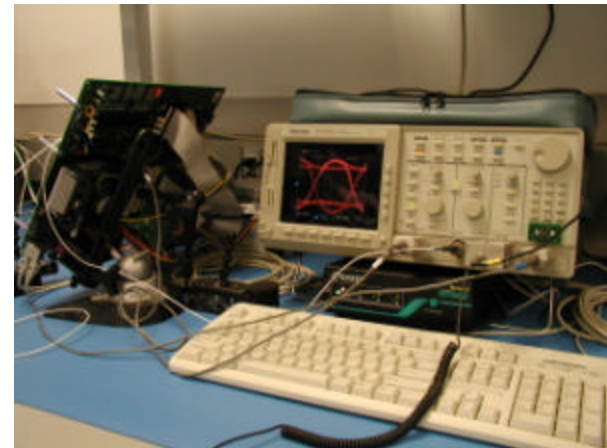
Frequency Domain



Time Domain



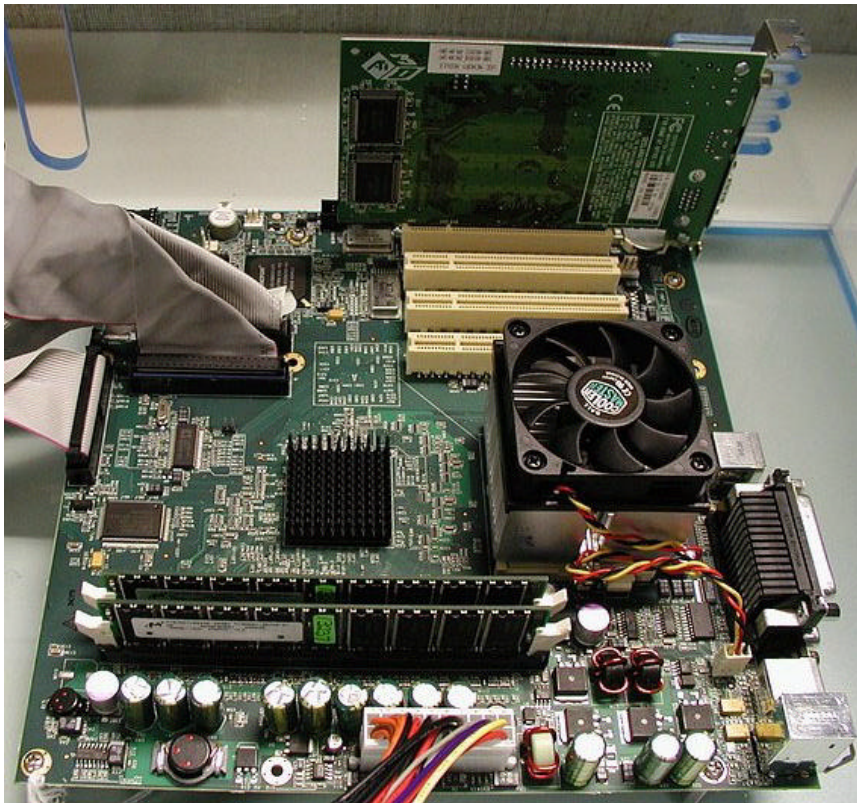
Reference Plane Analysis



Hardware Analysis

From Vision to Reality

Micron DDR Technology Hardware Demos



- **DDR333 Server**
 - 4-RDIMM system
- **DDR400 Desktop**
 - 2-UDIMM system
- **DDR II Desktop**
 - 533 Mb/s
 - 2-UDIMM system
 - Signal analysis demonstration

From Vision to Reality

Summary

- **End-use applications for memory have changed**
- **Diverse applications have different types of memory performance needs**
- **Memory innovation is required**
- **System designs have become more challenging**
- **Close collaboration of the system designer and memory supplier is necessary**
- **Enjoy VTF!**